COMPRESSION-BASED SEMANTIC-SENSITIVE IMAGE SEGMENTATION: PRDC-SSIS

Masahiro Nakajima, Toshinori Watanabe, and Hisashi Koga

Graduate School of Information Systems, University of Electro-Communications
1-5-1 Chofu-ga-oka, Chofu-shi, Tokyo 182-8585, Japan
nakajima@sd.is.uec.ac.jp, komkumei@gmail.com, koga@is.uec.ac.jp

ABSTRACT

This paper proposes PRDC-SSIS, a new compressibility-feature based semantic-sensitive image segmentation method using PRDC. One of the drawbacks of traditional signal (pixel-color) based image segmentation is the poor capability to capture the semantical information contained in the images. Because the semantic information tends to be carried by a set of neighboring pixels, rather than an individual pixel, we divide the image into patches and classify the patches based on their semantical contents. The crucial problem is classifying the patches into groups of similar patches according to their contents, and so we exploit the compressibility feature vector space of PRDC to accomplish this. An application of this method to an EO-image confirmed the proposed scheme can be carried out without any of the human-tailored target object models required by almost all traditional methods.

Index Terms— Image segmentation, semantic-sensitivity, compressibility feature

1. INTRODUCTION

This paper introduces a new semantic-sensitive image segmentation (SSIS) method, called PRDC-SSIS, which is based on the pattern representation scheme using data compression (PRDC). PRDC, previously proposed by one of the authors, is an improvement over traditional signal-based image segmentation [**?**][**?**][**?**]. One of the drawbacks of the traditional signal (pixel-color) based image segmentation prevalent in EO-image analysis is the poor capability to capture the semantical information contained in the images [**?**][**?**][**?**]. As examples, we introduce the simple artificial images in Fig. **??**. Our brain vision systems can easily find the semantical objects: sea, coastline, road, green belt, river, farm, house, and so on. However, using the traditional image segmentation methods, we only obtain meaningless similarly colored pixel segments, and we do not grasp the semantical objects described above. Different objects are misjudged as similar (Fig. **??** (a), (b)), or a compound object with different color components is not recognized (c). Now, let us divide the image into a set of patches, as shown in Fig. **??**. As each of the patches tends to contain various components, we can assign a meaningful semantical label to the patches, as follows: the gray road and apartments are judged to be different due to different patch contents (a, b) and the bare (brown) and green ground are judged to comprise a farm (c).

Based on these observations, we introduce a new semantic-sensitive image segmentation method that applies the following conditions.

(F1) The given image is divided into patches.

(F2) All the patches are classified into several semantically similar groups

For F1, the appropriate patch size varies depends on the target object. If the patch size is too small, the patch is reduced to a pixel and we cannot get the semantic-sensitive segmentation. If the patch is too large, small objects cannot be extracted. For F2, we must classify a tremendous variety of patch contents, particularly in EO-images. This can be observed even in Fig. **??**. Patches that seem to belong to the same class may have contents varying in color, shape, and configuration. So, we need a patch classification method that is very robust to the variations in contents.

The PRDC-SSIS solution is as follows.

(SF1) For F1, we give a set of image segmentation results for several patch sizes provided by the user. For the small patch sizes, small target objects are captured, whereas for the large patch sizes, large target objects are captured. We adopt this user-defined method because the sizes of objects and the image resolutions vary beyond the scope of computerization.

(SF2) For F2, we exploit the universal compressibility vector feature of PRDC that provides us the means to classify the patches with indefinite variety.

The compressibility of data, especially text, was initially proposed by Ziv and Lempel [**?**], and Welch [**?**]. Also proposed was an approximation of the conceptual Kolmogorov complexity $K(x)$ [**?**] that was a counter idea to the statistical Shannon entropy $H(X)$. The significant contribution of $K(x)$ is that it is applicable to individual data x, not to a statistical ensemble X. PRDC exploits the conditional compressibility to obtain a universal feature vector space for multimedia data analysis. Under the PRDC scheme, we can eliminate the laborious steps of statistical model building for varying patches that cannot be avoided in traditional methods.

978-1-4673-1159-5/12/$31.00 ©2012 IEEE

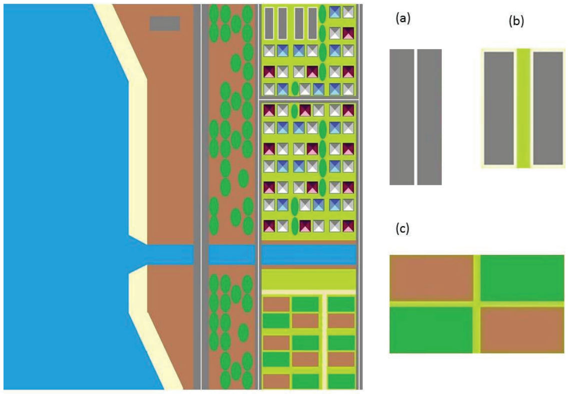

Fig. 1. Pixel-color based, semantic-insensitive segmentation (left); gray roads and apartments are judged as similar (a, b); bare and green ground are judged as two different objects (c)

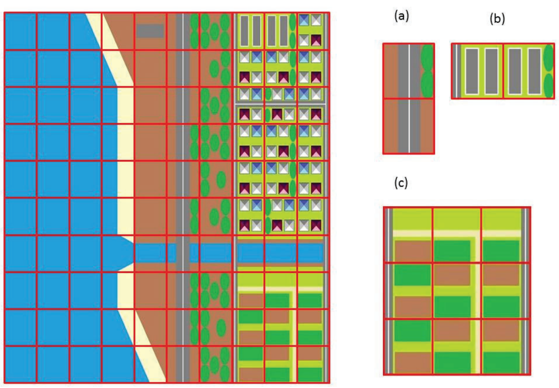

Fig. 2. Patch-based semantic-sensitive view (left); gray roads and apartments are judged to be different due to their contents (a, b); bare and green ground are judged as farm areas (c)

In this paper, Section **??** gives the design overview of PRDC-SSIS and Section **??** reports the experimental results.

2. OVERVIEW OF PRDC-SSIS

Fig. **??** shows an overview of PRDC-SSIS. In the following, the motivations and the implementations are briefly explained for each of the main functions.

2.1. Image division into patches

Following the discussions of SF1 above, we divide the image into square patches. Typically, users are expected to define the patch size that covers the target objects of concern. This step corresponds to A in the diagram.

2.2. Patched image segmentation and patch labeling

Step 1: Among the many possibilities, we use the minimum spanning tree (MST) based patched image segmentation. We introduce a graph in which each node represents a patch, and four edges connect the patch to the upper, right, lower, and left patches. Each edge carries a weight representing the attribute difference between the two end nodes. We use the mean color difference as the weight. Then, we create the MST of the graph, and, on the basis of a threshold, cut the large edges to obtain a set of subtrees. Each subtree represents a similarly colored patch group [**?**]. This step corresponds to B in the diagram.

Step 2: For each patch, we give a content-sensitive label. We apply the PRDC scheme to each patch as follows. We select the text of a patch by first getting its pixel-MST, traverse it by the similar weight first manner, and get the character at each arriving node by using both the reached node color and the reaching direction. In this way, we can encode both the color and local shape (color contour) information into the text. Out of all the text derived from all the patches, we sample several texts and LZW-compress them to obtain a set of dictionaries to span the compressibility feature space. Then we map all of the patch texts into this space by compressing the texts by using the dictionaries. We obtain a set of feature vectors of all the patches, and we classify the feature vectors by using, for example, the k-means method, and obtain their labels. As each text encodes both the pixel color and the local shape (contour running directions) of the patch contents, the output label reflects the patch contents [**?**][**?**][**?**]. This step corresponds to C in the diagram.

2.3. Patch grouping and group labeling

Finally, by following the PRDC process used in step 2 above, we have the patch groups and their labels. The difference is only the input texts. In this case, the text of a patch group is generated by tracing the patch label along the MST edge connecting the group. (The MST can be reused.) This step corresponds to D in the diagram. In an ideal case, when all the patches in a group have an identical patch label, the patch grouping step is not needed. However, this is rarely, if ever, the case and so this step contributes to getting large patch groups across patches with occasionally changed labels. Also, we set aside the noisy singleton texts (containing only one character representing a singleton patch) from the processing in C and return them to the most similar neighboring group of which mean compression vector nearest to the singleton.

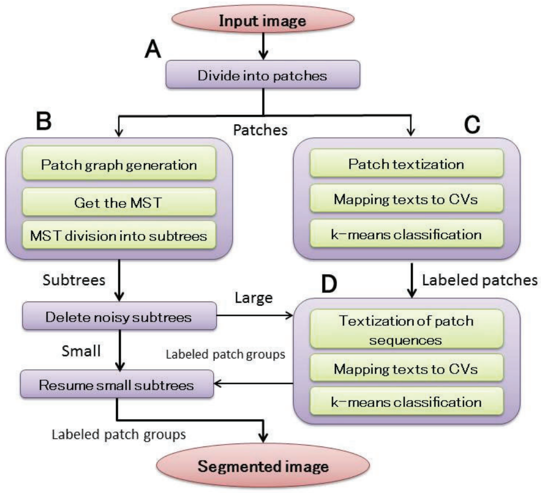

Fig. 3. The PRDC-SSIS diagram

3. EXPERIMENTS

3.1. Experimental design

Using the EO-image of size [1280,1024]-pixels in Fig. **??**, we compare the segmentation results under the traditional semantic-insensitive (color-based) method and the semantic-sensitive PRDC-SSIS method. For the PRDC-SSIS method, we give two results for two different patch sizes: [64,64]- and [128,128]-square patches. In addition, we give qualitative evaluations of the results. The main parameters used are as follows: the MST-cutting threshold in B of Fig. **??** is the top 25% point of all edge weights, and the compressibility feature space dimension and the k-means parameters in C of Fig. **??** are 40 and 20, respectively. These parameters are 20 and 15, respectively, in D. These are tuned by preset experiments.

3.2. Results and considerations

Fig. **??** shows the semantic-insensitive segmentation. As only the pixel color information is used, similarly colored segments tend to be identically labeled (not shown in the figure). Fig. **??** shows the semantic-sensitive segmentation under the patch size of [64,64]-pixels. Identical patch groups are represented by the identically colored patch squares. We observe that the following semantical objects can be segmented: sea (black), seashore (red), farming area (white), car parking (gray), golf course (yellow), and green area (blue). Fig. **??** shows the result for the patch size of [128,128]-pixels. Similar objects are extracted as above, such as sea (gray), seashore (blue), farming area and seashore resort (brown), car parking (gray), golf course (pink and blue), and green area (orange). But two different objects, the farming area (right bottom) and the seashore resort (middle top), are merged. The classification F-values, $2 \times recall \times relevance/(recall + relevance)$, of these semantic objects compared to human visual classification are sea (0.98), seashore (0.79), farming area (0.89), car parking (0.50), golf course (0.84), and green area (0.73). As shown by the results, we can conclude that the proposed PRDC-SSIS succeeded in extracting the semantical objects contained in the given EO-image. It should be noted that PRDC-SSIS extracted the semantical objects without using the human-tailored target object models inevitable in many traditional statistical image analyses.

Fig. 4. Seaside resort

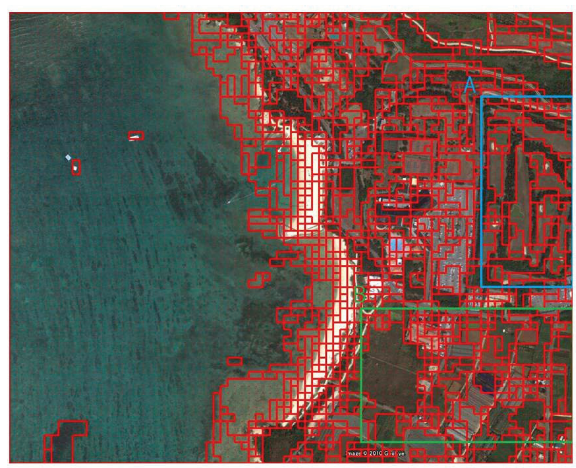

Fig. 5. Color based segmentation

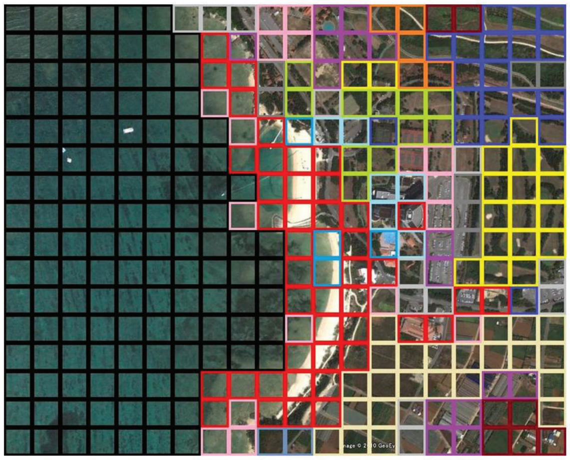

Fig. 6. Semantic-sensitive segmentation using [64,64]-pixel patches

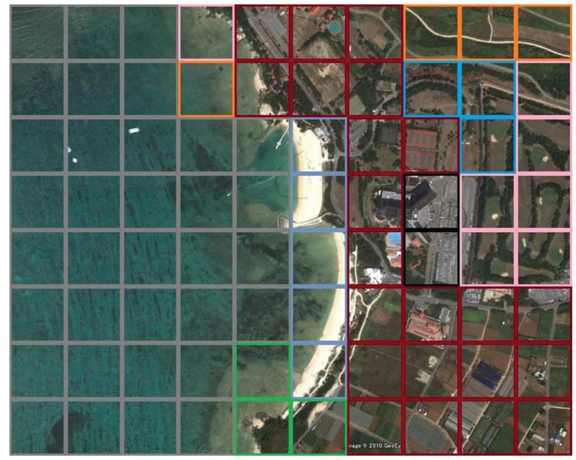

Fig. 7. Semantic-sensitive segmentation using [128,128]-pixel patches

4. CONCLUSIONS

This paper proved the capability of a new type of image segmentation using PRDC-SSIS.
(1) Instead of the traditional pixel-color based methods, the patches of PRDC-SSIS contain rich information, such as color, shape, configuration, etc., all of which suggest the existence of objects.
(2) To classify tremendously varying patches by their contents, the PRDC scheme is exploited to map patches into the compressibility feature vector space, wherein similar patches are classified and given labels.
(3) Similar patches are grouped and again their group labels are determined by using the PRDC scheme. The extracted groups are semantical objects in the given image.
An experimental application to an EO-image proved that PRDC-SSIS can output semantic-sensitive land cover segmentation completely free of human-tailored target object models.

5. ACKNOWLEDGEMENT

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan, Grant-in-Aid for Scientific Research (C), 22500122, 2010.

6. REFERENCES

[1] A. N. Kolmogorov, "Three approaches to the quantitative definition of information," *Prob. Inf. Transmission,* Vol. 1, pp. 1-7, 1965.

[2] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE trans. IT,* Vol. 22, No. 1, pp. 75-81, 1976.

[3] T. A. Welch, "A technique for high-performance data compression," *IEEE Computer*, Vol. 17, No. 6, pp. 8-19, 1984.

[4] T. Watanabe, K. Sugawara and H. Sugihara, "Map generation from aerial imagery using data compression," *proc. MVA 2000*, Vol. 17, No. 6, pp. 596-599, 2000.

[5] T. Watanabe, K. Sugawara and H. C. Park, "A new universal media featuring scheme using data compression," *Proc. Int'l Conf. on Media Futures*, pp. 261-264, 2001.

[6] T. Watanabe, K. Sugawara and H. Sugihara, "A new pattern representation scheme using data compression," *IEEE TPAMI*, Vol. 24, No. 5, pp. 579-590, 2002.

[7] C. Mathieu, I. Magnin and C. B. Porcher, "Optimal stochastic pyramid: segmentation of MRI data," *Proc. SPIE*, Vol. 1652, pp. 14-22, 1992.

[8] J. R. Lersch, A. E. Iverson, B. N. Webb and K. F. West, "Segmentation of multiband imagery using minimum spanning trees," *Proc. SPIE*, Vol. 2758, pp. 10-18, 1996.

[9] X. X. Zhang and Y. M. Yang, "Minimum spanning tree and color image segmentation," *IEEE ICNSC*, pp. 900-904, 2008.